

7 Summary

A porcine P1-derived artificial chromosome (PAC) library from a male German Landrace pig was constructed using pCYPAC2 vector and *E. coli* DH10B. A total of 90,240 clones were generated and individually transferred into microtiter plates. To determine the average insert size, 150 clones were randomly selected and separated by pulsed field gel electrophoresis. With an average insert size of 119.1 kb, the PAC library covers approximately 3.24 genome equivalents. The stability of nine clones was followed through 110 generations showing no reduction of the insert size. To test the probability of identifying a specific region within the library, the library was screened for the presence of 7 type I and 5 type II loci. Most of the loci (10/12) were present in the library more than once. To determine the percentage of chimerism within the PAC library, 6 clones were analyzed using *in situ* hybridization (FISH) on metaphase chromosome. With this test, three previously described type II loci which have not yet been physically assigned were mapped to chromosome 6p1.1-q1.2 (SW855), 6q1.2-q2.1 (S0300), and 6q1.2-2.12 (SW1129). Two isolated PAC clones designated IVMP714B01415 and IVMP714B05121 were assigned to chromosome 13q3.1-q3.2 and 11p1.3-p1.4, respectively. The location of IVMP714B01415 was also confirmed by screening a somatic hybrid cell panel. Finally a PAC harbouring the porcine triadin gene was assigned to chromosome 1q1.3-q1.4.

The porcine PAC library presented in this thesis will support the efforts and advance the possibilities for genome analysis, contig cloning, gene isolation and suggest that PACs are ideal building blocks for high resolution physical maps. This will ultimately help to identify and diagnose inherited disorders in swine. In order to share the clones of the library with other research groups, the library has been included in the Reference Library System (No. 714) of the resources Center/Primary Database of the German Human Genome Project.

Hiam AL- Bayati

Konstruktion und Charakterisierung einer porcinen PAC-Genbank mit 3.2 Genomäquivalenten sowie cytogenetische Lokalisation von 6 Typ I und Typ II Genorten

8 Zusammenfassung

Die Konstruktion von Genbanken mit großen Inserts ist eine wesentliche Voraussetzung für die Erstellung von physikalischen Karten und funktionellen Analysen komplexer Genome. Während der Kartierungsprozeß für das menschliche Genom rapide Fortschritte macht, blieb er für landwirtschaftliche Nutztiere dahinter zurück. Der Grund dafür liegt darin, daß für den Menschen zahlreiche Genbanken für die Klonierung in großem Maßstab zur Verfügung stehen, für Haustiere jedoch erst wenige. Bis jetzt wurden vom Schwein vier Hefe- und eine bakterielle künstliche Chromosomenbank kloniert. Eine von mehreren Möglichkeiten, diese Entwicklung zu beschleunigen, ist die Erstellung von Genbanken mit 6- bis 10 fachen Genomäquivalenten (McPherson, 1997). Grundsätzlich erfordern die physikalische Kartierung und das „chromosome walking“ Genbanken, die schnell zu erzeugen, leicht zu durchmustern und zu handhaben sind, und die frei von chimären Klonen sind. Zu diesem Zweck wurden verschiedene Klonierungssysteme entwickelt, die jeweils Vor- und Nachteile haben.

Der Umgang mit Bakteriophagen und Cosmiden ist in den meisten Fällen einfach, jedoch machen die Insertgrößen dieser Vektorsysteme (zwischen 24 und 45 kb) ihren

Gebrauch kompliziert und schränken ihre Anwendung für die Erstellung von „long-range“ Genkartierungen aus. Außerdem neigen Cosmide zur Rekombination (Yokobata *et al.*, 1990, Kim *et al.*, 1992).

Künstliche Hefechromosomen (Yeast Artificial Chromosome, YAC) wurden von Burke u. a. entwickelt (Burke *et al.* 1987). Der Hauptvorteil von YACs ist ihre Fähigkeit, Inserts von mehr als einer Megabase aufzunehmen (Larin *et al.*, 1991, Dausset *et al.*, 1992). Große YACs haben zwar im Vergleich mit kleineren Insertklonen für die Entwicklung physikalischer Karten Vorteile, es wurden jedoch auch Probleme bei der Konstruktion von YAC-Genbanken beschrieben, z.B. die geringe Klonierungseffizienz, Schwierigkeiten beim Trennen und Isolieren der YAC-DNA von der restlichen Hefe-DNA, die genomische Instabilität und der hohe Anteil an chimären Klonen (Smith *et al.*, 1990, Neil *et al.*, 1990, Green *et al.*, 1991, Larionov *et al.*, 1994, Sternberg, 1994).

Die Nachteile von YACs und Cosmiden wurden durch die Entwicklung des Bakteriophagen P1-Systems umgangen (Sternberg *et al.*, 1990; Pierce *et al.*, 1992; Rao *et al.*, 1992). Dieses System hat wünschenswerte Eigenschaften wie Stabilität und gute Trennung der Rekombinanten von Nicht-Rekombinanten. Allerdings beträgt die Klonierungskapazität des P1-Systems nur rund 95 kb, und die komplizierten Anforderungen bei der *in vitro*-Verpackung schränken seine Brauchbarkeit ein.

Ein neues Klonierungssystem, bakterielle künstliche Chromosomen (BAC), wurde von Shizuya u. a. entwickelt (Shizuya *et al.*, 1992). Dieses System basiert auf dem *E. coli* F-Faktor-Plasmid mit einer Kapazität bis zu 300 kb. Den Vorteilen des BAC-Systems wie Stabilität, hohe Effizienz, relativ geringe Chimärenbildung, steht allerdings die geringe Ausbeute an DNA entgegen, die aus einer Kolonie gewonnen werden kann (Koppes, 1992; Shizuya *et al.*, 1992).

In jüngster Zeit haben Ioannou und Mitarbeiter (Ioannou *et al.*, 1994) ein System entwickelt, das die wünschenswerten Eigenschaften des P1-Klonierungs-Systems mit denen des BAC-Systems kombiniert, die sog. P1-derived artificial chromosome (PAC).

Dieses System ist benutzt worden, um einige partielle und vollständige PAC-Genbanken von Mensch, Ratte, Hund, Pavian und von Mikroorganismen zu konstruieren (Ioannou *et al.*, 1996; Gingrich *et al.*, 1996; Woon *et al.*, 1998; Osoegawa *et al.*, 1998; Piper *et al.*, 1998). Das PAC-Klonierungssystem, mit dem Fragmente von 100 bis 300 kb in bakteriellen Zellen kloniert werden können, zeichnet sich durch hohe Stabilität, geringe bis gar keine Chimärenbildung, hohe Klonierungseffizienz, hohe Ausbeute von DNA aus und ermöglicht die Erstellung physikalischer Kartierungen von hoher Auflösung (Ioannou *et al.*, 1995; Ashworth *et al.*, 1995; Nothwang *et al.*, 1997; Hubert *et al.*, 1997; Matsumoto *et al.*, 1997; Nechiporuk *et al.*, 1997).

Alle bisherigen Klonierungsmethoden haben Stärken und Schwächen. Daher wurde die vorliegende porcine PAC-Genbank auf der Basis entsprechender Erfahrungen erstellt. Für die Erstellung der Genbank wurde pCYPAC2 als Vektor verwendet, da dieser Vektor Segmente verschiedener Sequenzen enthält, die das Klonierungssystem sehr effizient machen. Hierzu gehört z.B. das SacBII-Gen, das bei der Unterdrückung nicht-rekombinanter Klone eine wichtige Rolle spielt (Shizuya *et al.*, 1992).

Für die Erstellung der Genbank wurden zwei DNA-Quellen benutzt und partiell mit *Mbo*I gespalten. Die Größen-selektierte DNA wurde aus PFGE-Gelen geschnitten und mit gespaltenem dephosphorylierten pCYPAC2-Vektor ligiert

Die Ligationen wurden in *E.coli* DH10B elektroponiert. Insgesamt wurden 90.240 Klone erzeugt und einzeln in Mikrotiterplatten übertragen. Zur Bestimmung der durchschnittlichen Insertgröße wurden 150 Klone zufällig ausgewählt und durch Pulsfeldgelelektrophorese getrennt. Die Größe variierte zwischen 35 und 272 kb mit einem Mittelwert von 119,1 kb. Zieht man die korrigierte Zahl der Rekombinanten und die durchschnittliche Insertgröße in Betracht, so enthält die PAC-Genbank ca. 3.24 Genomäquivalente. Daraus ergibt sich eine Wahrscheinlichkeit von 96,1% für die Anwesenheit einer spezifischen DNA-Region in der Genbank. Die Stabilität von neun PAC-Klonen mit verschiedenen Insertgrößen wurde über 100 Generationen als stabil getestet. Um die Empfindlichkeit der Detektion zu erhöhen, wurde zusätzlich eine Hybridisierung mit einer SINE-Sonde durchgeführt. Die Analyse bestätigte die Ergebnisse der vorangegangenen Beobachtungen. Das Fehlen eines Signals in einem Klon ist sehr wahrscheinlich darauf zurückzuführen, daß sich in dem Insert dieses Klons kein SINE befand. Um eine effiziente Durchmusterung auf der Basis von einer PCR zu ermöglichen, wurde die DNA der Genbank vereinigt. Das DNA-Pooling-System war - bis auf geringfügige Modifikationen - im wesentlichen dasselbe wie von Kim *et al.* (1986) beschrieben. Die anfängliche Anzahl von PCRs, die für eine Genbankdurchmusterung durchgeführt werden muß, wurde durch die Zusammenfassung von fünf Superpools zu einem Super-Superpool reduziert. Die Ergebnisse zeigen, daß das eindimensionale Poolingsystem weniger kompliziert ist als etwa das von Asakawa *et al.* (1987) beschriebene. Es erlaubt die Identifizierung eines einzigen positiven PAC in vier aufeinanderfolgenden PCR-Screening.

Die Genbank wurde auf die Anwesenheit von sieben Typ I (Calcitonin-Rezeptor, β -1,2-N-Acetyl-glucos-aminyl-transferase II, Phospho-gluconat-Dehydrogenase, Proteolipid Protein, Ryanodin-Rezeptor 1, Secretory carrier membrane protein 1, Triadin) und fünf Typ II Genorten (DGC626, SW4, SO300, SW855, SW1129) untersucht. Die meisten (10/12) Genorte konnten mehrfach in der Genbank identifiziert werden.

Der Anteil chimärer Klone in der PAC-Genbank wurde bestimmt. Die Lokalisation von einem PAC-Klon (IVMP714BO5415) wurde durch Durchmusterung eines porcinen somatischer Hybridzellen-panels analysiert. Außerdem wurde mit diesem PAC-Klon und fünf weiteren eine FISH-Analyse durchgeführt. Nur bei einem Klon wurden wiederholt Signale auf zwei Chromosomen detektiert, die sich jedoch auf eine Kontamination des Klons zurückführen ließen. Das Vorkommen mehrerer Signale wurde auch von Kim *et al.* (1996) beschrieben.

Mittels FISH-Analyse wurden drei bereits früher beschriebene, aber noch nicht physikalisch zugeordnete Typ-II-Genmarker auf Chromosom 6p1.1 - q1.2 (SW855), 6q1.2 - q2.1 (SO300) bzw. 6q1.2-q2.1 (SW1129) lokalisiert. Diese Genorte entsprechen den Kartierungspositionen 77.2 cM, 78.4 cM bzw. 80.2 cM.

Zwei ausgewählte PACs wurden dem Chromosom 13q3.1-q3.2 und 11p1.3-p1.4 zugeordnet. Schließlich wurde ein PAC mit dem porcinen Triadin-Gen auf Chromosom 1q1.3-q1.4 lokalisiert. Die chromosomale Zuordnung der untersuchten Genorte liefert zusätzliche Anhaltspunkte für die vergleichende Genkartierung zwischen Mensch, Maus und anderen Spezies.

Die in dieser Arbeit erstellte porcine PAC-Genbank wird die Verfügbarkeit großer DNA-Fragmente für die Genomanalyse des Schweines erhöhen. Die Genbank wurde in das „Reference Library“ System (Nr. 714) des Deutschen Humanen Genomprojektes aufgenommen.